

Brain Age Prediction in Generalized Anxiety Disorder using a Convolutional Neural Network

Corey Richier

`corey.richier@nih.gov`

University of Illinois at Urbana-Champaign

André Zugman

<https://orcid.org/0000-0002-6126-7759>

Anita Harrewijn

Elise Cardinale

Parmis Khosravi

Moji Aghajani

Leiden University <https://orcid.org/0000-0003-2040-4881>

Willem Bruin

<https://orcid.org/0000-0002-0672-8903>

Kevin Hilbert

<https://orcid.org/0000-0002-7986-4113>

Narcís Cardoner

Institut d'Investigació Biomèdica Sant Pau (IIB-Sant Pau), Hospital de la Santa Creu i Sant Pau. Centro de Investigación Biomédica en Red de Salud Mental, Instituto de Sa <https://orcid.org/0000-0001-9633-0888>

Daniel Porta Casteràs

Hospital Sant Pau <https://orcid.org/0000-0002-3921-3607>

Marta Cano

Institut de Recerca Sant Pau (IR SANT PAU) <https://orcid.org/0000-0003-0675-9483>

Savannah Gosnell

Ramiro Salas

Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston TX, USA <https://orcid.org/0000-0002-1105-566X>

Andrea Jackowski

Federal University of São Paulo <https://orcid.org/0000-0001-8842-5406>

Pedro Pan

<https://orcid.org/0000-0002-1943-6520>

Giovanni Salum

Universidade Federal do Rio Grande do Sul

Karina Blair

Boys Town National Research Hospital

James Blair

Boys Town National Research Hospital <https://orcid.org/0000-0002-6377-2361>

Mohammed Milad

Katie Burkhouse

K. Luan Phan

The Ohio State University

Heidi Schroeder

Jeffrey Strawn

Katja Beesdo-Baum

Neda Jahanshad

University of Southern California

Sophia Thomopoulos

University of Southern California

Jared Nielsen

Neuroscience Center, Brigham Young University <https://orcid.org/0000-0002-2717-193X>

Jordan Smoller

Jair Soares

McGovern Medical School <https://orcid.org/0000-0002-5466-5628>

Benson Mwangi

University of Texas Health Science Center

Mon-Ju Wu

University of Texas Health Science Center <https://orcid.org/0000-0002-0540-8027>

Giovana Zunta Soares

University of Texas Health Science Center

Michal Assaf

Olin Neuropsychiatry Research Center <https://orcid.org/0000-0001-5307-1986>

Gretchen Diefenbach

Paolo Brambilla

University of Milan <https://orcid.org/0000-0002-4021-8456>

Eleonora Maggioni

Politecnico di Milano <https://orcid.org/0000-0003-4331-4029>

David Hofmann

University of Münster <https://orcid.org/0000-0003-3494-634X>

Thomas Straube

University of Münster

Carmen Andreescu

University of Pittsburgh School of Medicine <https://orcid.org/0000-0003-3767-5127>

Rebecca Price

Gisele Manfro

<https://orcid.org/0000-0002-0109-7591>

Federica Agosta

IRCCS San Raffaele Scientific Institute <https://orcid.org/0000-0003-3121-4979>

Elisa Canu

San Raffaele Scientific Institute, Vita-Salute San Raffaele University

Camilla Cividini

Massimo Filippi

IRCCS San Raffaele Scientific Institute <https://orcid.org/0000-0002-5485-0479>

Milutin Kostic

Ana Munjiza Jovanovic

Brenda Benson

Gabrielle Freitag

Ellen Leibenluft

National Institute of Mental Health Intramural Research Program

Grace Ringlein

Kathryn Werwath

Hannah Zwiebel

Hans Grabe

University Medicine Greifswald

Sandra Van der Auwera

University Medicine Greifswald <https://orcid.org/0000-0002-1757-7768>

Katharina Wittfeld

University Medicine Greifswald <https://orcid.org/0000-0003-4383-5043>

Henry Völzke

University Medicine Greifswald

Robin Bülow

University Medicine Greifswald <https://orcid.org/0000-0003-1884-5784>

Nicholas Balderston

University of Pennsylvania <https://orcid.org/0000-0002-8565-1544>

Monique Ernst

Lilianne Mujica-Parodi

Helena van Nieuwenhuizen

Hugo Critchley

University of Sussex <https://orcid.org/0000-0002-2445-9284>

Elena Makovac

Matteo Mancini

Frances Meeten

Crisitina Ottaviani

Sapienza University of Rome <https://orcid.org/0000-0002-5240-4387>

Gregory Fonzo

The University of Texas at Austin Dell Medical School <https://orcid.org/0000-0002-0213-1034>

Martin Paulus

Laureate Institute for Brain Research, Tulsa, USA <https://orcid.org/0000-0002-0825-3606>

Murray Stein

University of California San Diego <https://orcid.org/0000-0001-9564-2871>

Raquel Gur

University of Pennsylvania

Ruben Gur

University of Pennsylvania <https://orcid.org/0000-0002-4082-8502>

Antonia Kaczurkin

Vanderbilt University <https://orcid.org/0000-0002-0943-3094>

Bart Larsen

University of Minnesota <https://orcid.org/0000-0001-7896-4342>

Ted Satterthwaite

Jennifer Harper

Michael Perino

Chad Sylvester

Qiongru Yu

Patrick McClure

Francisco Pereira

Ulrike Lueken

<https://orcid.org/0000-0003-1564-4012>

Dick Veltman

Functional Imaging Center and Department of Psychiatry, Vrije Universiteit Medical Center, DeBoelelaan 1117, Amsterdam, The Netherlands

Paul Thompson

Imaging Genetics Center, Mark and Mary Stevens Neuroimaging & Informatics Institute, Keck School of Medicine of the University of Southern California, Marina del Rey, California, USA

Nynke Groenewold

University of Cape Town <https://orcid.org/0000-0002-0865-8427>

Janna Marie Bas-Hoogendam

Leiden University <https://orcid.org/0000-0001-8982-1670>

Dan Stein

University of Cape Town and Neuroscience Institute <https://orcid.org/0000-0001-7218-7810>

Nic van der Wee

Leiden University Medical Center

Anderson Winkler

Daniel (S) Pine

NIMH, Bethesda, MD

Article

Keywords:

Posted Date: July 1st, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-6866544/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: **Yes** PMP has received speakers' honoraria from Daiichi Sankyo, Libbs, Sandoz, Novartis, and Abbot. NJ is MPI of a research related grant from Biogen Inc, for work unrelated to this manuscript. JCS is a consultant for PFIZER, SUNOVIAN, SANOFI, JOHNSON & JOHNSON, Livanova, and Boehringer Ingelheim. Material support (for IOL site, led by MA and GJD) was provided by Neuronetics. FA is Section Editor of NeuroImage: Clinical; has received speaker honoraria from Philips, Novartis, and Biogen Idec; and receives or has received research supports from the Italian Ministry of Health, AriSLA (Fondazione Italiana di Ricerca per la SLA), and the European Research Council. EC has received research supports from the Italian Ministry of Health. MF is Editor-in-Chief of the Journal of Neurology; received compensation for consulting services and/or speaking activities from Biogen Idec, Merck-Serono, Novartis, and Teva Pharmaceutical Industries; and receives research support from Biogen Idec, Merck-Serono, Novartis, Teva Pharmaceutical Industries, Roche, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and ARiSLA (Fondazione Italiana di Ricerca per la SLA). MF is Editor-in-Chief of the Journal of Neurology; received compensation for consulting services and/or speaking activities from Biogen Idec, Merck-Serono, Novartis, and Teva Pharmaceutical Industries; and receives research support from Biogen Idec, Merck-Serono, Novartis, Teva Pharmaceutical Industries, Roche, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and ARiSLA (Fondazione Italiana di Ricerca per la SLA). GAF owns equity in Alto Neuroscience and has served as a consultant for SynpaseBio AI and Alto Neuroscience. MPP is an advisor to Spring Care, Inc., a behavioral health startup, he has received royalties for an article about methamphetamine in UpToDate. MBS has in the past 3 years been a consultant for Actelion, Acadia Pharmaceuticals, Aptinyx, Bionomics, BioXcel Therapeutics, Clexio, EmpowerPharm, Epivario, GW Pharmaceuticals, Janssen, Jazz Pharmaceuticals, Roche/Genentech, and Oxeia Biopharmaceuticals; and has stock options in Oxeia Biopharmaceuticals and Epivario. DJS has received consultancy honoraria from Discovery Vitality, Johnson & Johnson, Kanna, L'Oreal, Lundbeck, Orion, Sanofi, Servier, Takeda and Vistagen. K.H. is a scientific advisor and received virtual stock options of Mental Tech GmbH, which develops an AI-based chatbot providing mental health support.

Brain Age Prediction in Generalized Anxiety Disorder using a Convolutional Neural Network

Corey Richier MS^{1,2*}, André Zugman PhD², Anita Harrewijn PhD^{2,3}, Elise M. Cardinale PhD^{2,4},
 Parmis Khosravi PhD², Moji Aghajani PhD⁵, Willem B. Bruin, PhD⁵, Kevin Hilbert PhD⁶, Narcis
 Cardoner MD, PhD^{7,8,9}, Daniel Porta-Casteràs PhD⁷, Marta Cano PhD^{7,8,9}, Savannah Gosnell PhD¹⁰,
 Ramiro Salas PhD¹⁰, Andrea P. Jackowski MD, PhD¹¹, Pedro M. Pan MD, PhD¹¹, Giovanni A. Salum
 MD, PhD¹², Karina S. Blair PhD¹³, James R. Blair PhD¹⁴, Mohammed R. Milad PhD¹⁶, Katie L.
 Burkhouse PhD¹⁷, K. Luan Phan MD¹⁸, Heidi K. Schroeder BS¹⁹, Jeffrey R. Strawn MD¹⁹, Katja
 Beesdo-Baum PhD²⁰, Neda Jahanshad PhD²¹, Sophia I. Thomopoulos BA²¹, Jared A. Nielsen PhD^{22,23,24},
 Jordan W. Smoller MD²³, Jair C. Soares MD, PhD²⁵, Benson Mwangi PhD²⁵, Mon-Ju Wu PhD²⁵,
 Giovana B. Zunta-Soares MD²⁵, Michal Assaf MD^{26,27}, Gretchen J. Diefenbach PhD^{27,28}, Paolo
 Brambilla MD, PhD^{29,30}, Eleonora Maggioni PhD²⁹, David Hofmann MSc³¹, Thomas Straube PhD³¹,
 Carmen Andreescu MD³², Rebecca B. Price PhD^{32,33}, Gisele G. Manfro MD, PhD³⁴, Federica Agosta
 MD, PhD^{35,36}, Elisa Canu PhD³⁵, Camilla Cividini MSc^{35,36}, Massimo Filippi MD^{35,36,37,38,39}, Milutin
 Kostić MD, PhD^{40,41}, Ana Munjiza Jovanovic MD, PhD⁴⁰, Brenda Benson PhD², Gabrielle F. Freitag
 MS², Ellen Leibenluft, MD², Grace V. Ringlein^{2, 42}, Kathryn Werwath BS⁴³, Hannah Zwiebel MD²,
 Hans J. Grabe MD^{44,45}, Sandra Van der Auwera PhD^{44,45}, Katharina Wittfeld PhD⁴⁴, Henry Völzke
 MD⁴⁶, Robin Bülow MD⁴⁷, Nicholas L. Balderston PhD⁴⁸, Monique Ernst MD, PhD⁴⁹, Lilianne R.
 Mujica-Parodi PhD⁵⁰, Helena van Nieuwenhuizen BE, BS⁵¹, Hugo D. Critchley DPhil⁵², Elena Makovac
 PhD⁵³, Matteo Mancini PhD⁵⁴, Frances Meeten DPhil, DClinPsych⁵⁵, Cristina Ottaviani PsyD, PhD^{56,57},
 Gregory A. Fonzo PhD⁵⁸, Martin P. Paulus MD⁵⁹, Murray B. Stein MD, MPH⁶⁰, Raquel E. Gur MD,
 PhD⁶¹, Ruben C. Gur PhD⁶¹, Antonia N. Kaczurkin PhD⁶², Bart Larsen PhD^{61,63}, Theodore D.
 Satterthwaite MD⁶¹, Jennifer Harper⁶⁴, Michael T. Perino PhD⁶⁴, Chad M. Sylvester MD, PhD⁶⁴,

25 Qiongru Yu⁶⁴, Patrick McClure PhD⁶⁵, Francisco Pereira PhD⁶⁵, Ulrike Lueken PhD⁶⁶, Dick J. Veltman
 26 MD, PhD⁶⁷, Paul M. Thompson PhD²¹, Nynke A. Groenewold PhD⁶⁸, Janna Marie Bas-Hoogendam
 27 PhD^{69,70,71}, Dan J. Stein MD, PhD^{72,68}, Nic J. A. Van der Wee MD, PhD^{69,71}, Anderson M. Winkler MD,
 28 DPhil^{2, 73}, Daniel S. Pine MD², and Chelsea K. Sawyers PhD^{2, 74}

29

- 30 1. Department of Psychology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
- 31 2. Emotion and Development Branch, National Institute of Mental Health, Bethesda, MD, USA
- 32 3. Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam,
- 33 The Netherlands
- 34 4. Department of Psychology, The Catholic University of America
- 35 5. Leiden University, Institute of Education & Child Studies, Section Forensic Family & Youth Care
- 36 6. Department of Psychology, HMU Health and Medical University Erfurt, Erfurt, Germany
- 37 7. Sant Pau Mental Health Research Group, Institut de Recerca Sant Pau (IR SANT PAU), Barcelona,
- 38 Spain
- 39 8. Centro de Investigación Biomédica En Red en Salud Mental (CIBERSAM), Instituto de Salud
- 40 Carlos III, Madrid, Spain
- 41 9. Instituto de Salud Carlos III (ISCIII), Madrid, Spain
- 42 10. Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine,
- 43 Houston, TX, USA
- 44 11. LiNC, Department of Psychiatry, Federal University of São Paulo, São Paulo, Brazil
- 45 12. Section on Negative Affect and Social Processes, Hospital de Clínicas de Porto Alegre,
- 46 Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
- 47 13. Center for Neurobehavioral Research, Boys Town National Research Hospital, Boys Town, NE,
- 48 USA
- 49 14. Child and Adolescent Mental Health Center, Copenhagen University Hospital – Mental Health
- 50 Services CPH, Copenhagen, Denmark
- 51 15. Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of
- 52 Copenhagen, Denmark
- 53 16. Department of Psychiatry, NYU School of Medicine, New York University, New York, NY, USA
- 54 17. Department of Psychology, Penn State University, University Park, PA, USA
- 55 18. Department of Psychiatry and Behavioral Health, The Ohio State University, Columbus, OH, USA
- 56 19. Department of Psychiatry & Behavioral Neuroscience, University of Cincinnati, Cincinnati, OH,
- 57 USA
- 58 20. Behavioral Epidemiology, Institute of Clinical Psychology and Psychotherapy, Technische
- 59 Universität Dresden, Dresden, Germany
- 60 21. Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck
- 61 School of Medicine, University of Southern California, Marina del Rey, CA, USA

- 62 22. Center for Brain Science & Department of Psychology, Harvard University, Cambridge, MA, USA
63 23. Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA
64 24. Psychology Department & Neuroscience Center, Brigham Young University
65 25. Center Of Excellence On Mood Disorders, Louis A. Faillace, MD, Department of Psychiatry and
66 Behavioral Sciences, The University of Texas Health Science Center at Houston, Houston, TX, USA
67 26. Olin Neuropsychiatry Research Center, Institute of Living, Hartford Hospital, Hartford, CT, USA
68 27. Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA
69 28. Anxiety Disorders Center, Institute of Living, Hartford Hospital, Hartford, CT, USA
70 29. Department of Neurosciences and Mental Health, Fondazione IRCCS Ca' Granda Ospedale
71 Maggiore Policlinico, Milan, Italy
72 30. Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy
73 31. Institute of Medical Psychology and Systems Neuroscience, University of Muenster, Muenster,
74 Germany
75 32. Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA
76 33. Department Psychology, University of Pittsburgh, Pittsburgh, PA, USA
77 34. Anxiety Disorder Program, Hospital de Clínicas de Porto Alegre, Department of Psychiatry, Federal
78 University of Rio Grande do Sul, Porto Alegre, Brazil
79 35. Neuroimaging Research Unit, Institute of Experimental Neurology, Division of Neuroscience,
80 IRCCS San Raffaele Scientific Institute, Milan, Italy
81 36. Vita-Salute San Raffaele University, Milan, Italy
82 37. Neurology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy
83 38. Neurophysiology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy
84 39. Neurorehabilitation Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy
85 40. Institute of Mental Health, University of Belgrade, Belgrade, Serbia
86 41. Department of Psychiatry, School of Medicine, University of Belgrade, Belgrade, Serbia
87 42. Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD,
88 USA
89 43. Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA
90 44. Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald,
91 Germany
92 45. German Center for Neurodegenerative Diseases (DZNE), Site Rostock/Greifswald, Greifswald,
93 Germany
94 46. Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany
95 47. Institute for Diagnostic Radiology and Neuroradiology, University Medicine Greifswald,
96 Greifswald, Germany
97 48. Center for Neuromodulation in Depression and Stress, University of Pennsylvania, Philadelphia, PA,
98 USA
99 49. Section on Neurobiology of Fear and Anxiety, National Institute of Mental Health, Bethesda, MD,
100 USA
101 50. Department of Biomedical Engineering, Stony Brook University, Stony Brook, NY, USA

51. Department of Physics, Stony Brook University, Stony Brook, NY, USA
52. Department of Neuroscience, Brighton and Sussex Medical School, University of Sussex, Brighton, UK
53. Department of Psychology, Brunel University London
54. "Enrico Fermi" Research Center, Rome, Italy
55. Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
56. Department of Psychology, Sapienza University of Rome, Rome, Italy
57. IRCCS Santa Lucia Foundation
58. Department of Psychiatry and Behavioral Sciences, The University of Texas at Austin Dell Medical School, Austin, TX, USA
59. Laureate Institute for Brain Research, Tulsa, OK, USA
60. Department of Psychiatry, School of Medicine and Herbert Wertheim School of Public Health, University of California, San Diego, La Jolla, CA, USA
61. Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA
62. Department of Psychology, Vanderbilt University, Nashville, TN, USA
63. Department of Pediatrics, University of Minnesota, Minneapolis, MN, USA
64. Department of Psychiatry, Washington University, St. Louis, MO, USA
65. Machine Learning Core, National Institute of Mental Health, Bethesda, MD, USA
66. Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany; German Center for Mental Health (DZPG), partner site Berlin-Potsdam
67. Department of Psychiatry, Amsterdam UMC, location VUMC, Amsterdam, The Netherlands
68. Department of Psychiatry & Neuroscience Institute, University of Cape Town, Cape Town, South Africa
69. Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands
70. Department of Developmental and Educational Psychology, Institute of Psychology, Leiden University, Leiden, The Netherlands
71. Leiden Institute for Brain and Cognition, Leiden, The Netherlands
72. South African Medical Research Council Unit on Risk & Resilience in Mental Disorders, Department of Psychiatry & Neuroscience Institute, University of Cape Town, Cape Town, South Africa
73. University of Texas Rio Grande Valley, Brownsville, TX, USA
74. Division of Child and Adolescent Psychiatry, Department of Psychiatry, Virginia Commonwealth University

*Corresponding author: Corey Richier. University of Illinois at Urbana-Champaign Department of Psychology. Psychology Building, 603 E Daniel St. Champaign, IL 61820 Email: coreyjr2@illinois.edu Phone: 217-333-0631 Fax: 217-333-2169

140 Supported by National Institute of Mental Health Intramural Research Program project ZIA-
141 MH002781

142

143

144

145

146

147

148

149

150

151

152

153

154

Abstract

Higher predicted brain age difference has been associated with several psychiatric disorders. Generalized anxiety disorder (GAD) is associated with markers of accelerated aging. In this study, we determined brain predicted age difference (PAD) in individuals with GAD and healthy controls (HC) as well as group differences in PAD variability using voxel-wise structural MRI. The training dataset included 3,511 controls, and the testing dataset included 1,595 individuals with GAD and 4,552 HC from the ENIGMA-Anxiety GAD Working Group. A convolutional neural network model using four input modalities per subject and a model ensemble approach was used to predict brain age. The PAD was then calculated by subtracting chronological age. Model performance was consistent with other image-based brain age prediction models with similar accuracy across the training set (mean absolute error (MAE) = 2.95 years) and HC in the testing set (MAE = 2.94). We found no evidence of accelerated brain aging in individuals with GAD, though we did find evidence for greater variation in PAD for individuals with GAD (Levene's test: $W = 442.98, p < .001$) and evidence for greater variability in PAD of those with GAD over 25 years of age. No relationships between PAD and clinical or demographic measures were found. To conclude, using large training and testing samples, the study found no significant association between GAD and PAD, although individuals with GAD had greater heterogeneity in brain-predicted age.

Introduction

Generalized anxiety disorder (GAD) is one of the most prevalent forms of psychiatric disorders, having a lifetime prevalence of 6.2% [1-3]. GAD has a peak onset in late adolescence and early adulthood and high comorbidity rate [2-4]. In this study, we leverage available large imaging datasets to investigate whether GAD is associated with an individual's predicted "brain age". Brain age modeling uses neuroimaging data to estimate an individual's chronological age, with deviations between predicted and actual age—referred to as the predicted age difference (PAD)—increasingly regarded as potential biomarkers of atypical neurodevelopment or neurodegeneration[5, 6]. These brain aging estimates may proceed faster or slower than would be expected. A positive predicted age difference (PAD) may indicate accelerated maturation, or an "older" brain than expected given a chronological age; conversely, a negative PAD may indicate a "younger" brain. Such investigations would be particularly relevant for GAD as individuals with this condition appear to have structural, functional, and cognitive differences [7-10], many of which occurring at critical phases of brain development, such as late childhood and adolescence

Higher brain PAD has been associated with a wide range of physical and psychological conditions including cognitive impairment [11, 12], traumatic brain injury [13], schizophrenia [14, 15], childhood adversity [16], and major depressive disorder [17]. More broadly, psychiatric disorders are associated with advanced brain age [17-20]. However, findings related to brain age and anxiety disorders are inconsistent. For example, one study found no relationship between dimensional measures of anxiety in youth, but no significant PAD [21]; however another study found a positive PAD of 2.91 years in adults (N=220) with anxiety disorders, including GAD, panic disorder (PD), and social anxiety disorder (SAD; [18]. Interestingly, this finding was only significant after accounting for medication use. More recently, in a study of individuals with SAD and GAD, adolescents with more severe GAD

showed a higher PAD compared to controls, and adolescents with more severe SAD symptoms showed a lower PAD [22]. To date, no work has examined brain age in individuals with GAD as a distinct group.

Given unclear evidence regarding the association of brain age and anxiety disorders, we aim to evaluate PAD and assess its relation to the presence of GAD. To that end, we use a convolutional neural network (CNN) model to predict the age of each individual from their structural magnetic resonance images (MRI) and compute their PADs. CNNs are deep learning models for computer vision inspired by the brain's visual cortex [23], and are commonly used to examine brain age [24-27] as they are well suited for multidimensional data. These models capture spatial relations and features in their input data and are sensitive to subject-level spatial information. This can be a strength relative to univariate models (voxelwise, vertex-wise, or feature-based), as CNNs can extract additional structural information from the image, encoded in spatial patterns. The nonlinear nature of deep learning models (including CNN) allows for detection of complex relations between the input data and the subject's age. Moreover, in healthy individuals, CNN models produce more accurate age estimations than other modeling frameworks, such as linear models or other non-neural network machine learning methods [28].

Our study has two main hypotheses: first, that PAD would be greater in participants with GAD compared to healthy controls (HC); and that larger variance in PAD in GAD. Second, participants with GAD might have a greater prediction error by the CNN, thus leading to increased variance in the PAD for this group. Other brain age research has found greater variability in age predictions for clinical groups relative to individuals without psychopathology [17, 18, 29]. To assess the second hypothesis, we conduct a test to verify the equality of variance between subjects with and without GAD. This could reflect greater heterogeneity in individuals with GAD potentially derived from factors such as treatment and disease severity while controls would form a more homogeneous group. Likewise, this variance

might change as a factor of age [30, 31]. Therefore, we also hypothesize that variance in age prediction will depend on age. As secondary goals we investigate if PAD will be related to medication, comorbidity, and age-by-diagnosis interaction. To address these questions, we use structural brain MRI data from the Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) Anxiety Working Group [32, 33]. This was the largest study to date investigating structural alterations between GAD and controls, and found no association between GAD and brain morphometry in a well-powered mega-analysis of 28 sites [33]. The present study builds on that work to examine if “brain age” might reveal structural differences in GAD that were not detected in the mass univariate analysis of the previous work. As PAD is estimated from a multivariate (CNN) model, it may have more sensitivity than mass-univariate testing as done in the previous ENIGMA-Anxiety GAD work. Moreover, previous work on other cohorts have shown different results when using “brain age” estimated from features obtained from FreeSurfer [21]. Previous work on brain PAD both within other ENIGMA working groups [17] and outside of ENIGMA [21, 22] have used a variety of machine learning methods to build and train brain-age prediction models. Based on the model accuracy findings from Jonnson et al. [34], showing the superiority of a model combining different measures obtained by T1-weighted scans, we build a new voxelwise CNN model rather than via feature extraction (e.g., from FreeSurfer) to limit the loss of information.

Methods

Sample Selection

The present analysis is a pre-registered (<https://osf.io/4wtdy>) mega-analysis [35] of structural brain MRI data, from an international multi-site sample. Our training sample of 3,511 individuals includes several publicly available neuroimaging samples of both pediatric and adult healthy controls,

including the Human Connectome Project [36], Open Access Series of Imaging Studies 3 (OASIS-3; [37], Philadelphia Neurodevelopmental Cohort (PNC; [38], and the Pediatric Imaging, Neurocognition, and Genetics (PING) study [39]. The present analysis only used images from a single time point in studies with longitudinal data. OASIS-3 [37] images were randomly selected across participants across all three timepoints to provide a wider distribution of ages within the sample.

The clinical test sample consists of data from the ENIGMA Anxiety Working Group [35] with 6,147 subjects from 32 studies. This includes ABCD [40], CMI [41] and BHRC [42] as well as data provided from ENIGMA Anxiety participating sites. As above we only used a single time point for each study. In the case of ABCD and BHRC we only include the baseline data. The selection process for determining the final number of images used in the model is summarized in Supplemental Table 1, which illustrates the procedure for inclusion of participants. We excluded subjects due to failed processing, missing covariates (age, sex, diagnosis), comorbid psychosis, bipolar disorder or ASD, and subjects with no GAD diagnosis but any other psychiatric diagnosis.

Neuroimaging data preprocessing

All data were run through the FSL voxel-based morphometry (VBM) pipeline [43, 44], providing voxelwise output [45]. Such output is amenable for use with CNNs, while still capturing between-subject variability on brain morphology. Structural T1-weighted MRI brain images were first standardized and underwent intensity normalization. Afterwards, non-brain tissue was removed. Following this, the T1-weighted images were bias corrected. Images were then segmented into gray and white matter maps, modulated, and smoothed. It has been demonstrated that an increased number of imaging modalities results in better prediction [34]; thus, for this analysis we used four images derived from the VBM pipeline: grey matter, white matter, T1-weighted image, and the image of the Jacobian of the non-linear transformation used to map the individual T1-weighted image to the reference standard

space [46]; the latter provides a measure of stretches and shrinkages needed to non-linearly align each subject's image to a standard space, in this case, the MNI152 [47].

No data harmonization step was implemented (e.g., ComBat; [48]) due to recent demonstrations that not applying site/scanner harmonization produces lower mean absolute error (MAE) for brain age prediction than applying harmonization methods [49]. After the VBM process was completed, images were converted to NumPy arrays [50] to increase speed of image loading, and were cropped from their original dimension of 91 x 109 x 91 to 75 x 88 x 78 to reduce the amount of empty space in the image surrounding the brain tissue.

Quality control

Quality control procedures consisted of visual inspection to exclude any images that were misshapen or contained any other processing errors. Additionally, subjects who had errors in their covariate entries (see Supplemental Table 1) were excluded. After quality control, the remaining usable sample consisted of 8,063 HC and 1,595 participants with a diagnosis of GAD, generating a total sample size of 9,658. Tables 1 and 2 detail the number of participants included from each sample for the training and testing samples. Data use guidelines for all research sites were followed. Appropriate consent and assent procedures were obtained for all participants at all sites. Individuals with GAD had either a lifetime or present diagnosis of GAD. These diagnoses were established at each site via clinician-based assessment [32].

Train-test split procedure

For the test and train split procedure, the total number of subjects (9,658) was split into two groups: (a) subjects from samples including GAD cases and controls, and (b) subjects from samples composed of only controls. The latter was used as the training sample, which consisted of 3,511 individuals (Table 1). Ten percent of the training sample was used to track model performance while

training (referred in the results as *validation sample*). The test sample (n=6,147; Table 2) consisted of data from studies that contained subjects with GAD and controls.

Brain age model

The present analysis used the PyTorch library for the creation of the CNN model [51]. The input to the network was a 5D array consisting of batch size (the number of subjects given to the model at a time for training) by channels by X image dimension by Y image dimension by Z image dimension. The four channels were the gray and white matter segmentations, as well as the T1-weighted image and Jacobian images, all in MNI152 space.

The network architecture includes five repeated convolutional blocks, each containing 3D convolutional layers, batch normalization, ReLU activations and max pooling (Figure 1). This block was repeated five times, at each step doubling the number of convolutional filters but halving the size of the image via the max pooling layer. Thus, the number of filters was 8 for the first layer, and the final number of filters was 128. We used a 3x3x3 kernel with padding of size 1. Max pooling layers had a stride distance of 2 to downsize each dimension image by half. Layers following batch normalization utilized a ReLU activation function. After the last block, the image was flattened to a 1,024 size fully connected dense layer and then connected to two more fully connected layers of size 128 and 32 before connecting to the output node where age predictions were made with a ReLU function. Between each fully connected layer, a dropout layer was added with a rate of 50 percent, as is recommended for CNNs [52]. The total number of model parameters was 1,020,721.

The MAE was measured across each batch to track model performance and as the model loss function. The mean squared error (MSE) was also tracked but not used to optimize model performance. The Adaptive Moment Estimation (Adam) optimizer was used to update the model weights [53]. Gradient norm clipping was used to prevent exploding gradients [54], which occurs when the gradients

of the model parameters become excessively large during training, leading to instability. The maximum gradient value was set to 1 and learning rate was set to 0.0001.

A homogeneous ensemble strategy was used to combine the predictions of multiple models [55]. Other machine learning prediction work has used ensemble strategies with deep learning models [55, 56], including brain age prediction [57], which has been found to generate better predictive results. Since neural networks may converge to several possible local minima, it is ideal to use strategies that provide stable results. Instability in results would complicate downstream analysis and interpretation of results generated by a CNN. As is recommended when using homogeneous ensembles, it is best practice to introduce randomness into the constituent models, usually in the form of a diversified training set [55]. An ensemble of eight model instantiations was created for the present analysis. Each instantiation was initialized with random weights, but the architecture for each was identical. Ten percent of the training data was randomly selected to serve as the validation data for each model, such that each model in the ensemble had a different set of 90% of the full training sample. After the training process, the eight models' age predictions for each subject in the test set were then averaged together. These averaged predictions were treated as the overall predicted ages to be used in subsequent analyses and subtracted from the subject's true chronological age to generate the overall PAD for all participants within the entire test set.

Statistical Analysis

All statistical tests were conducted in PALM - Permutation Analysis of Linear Models [58]. The PAD measure was entered as the dependent variable for all models. Three general linear models (GLMs) were tested (Supplementary Table 3): one was a simple model with only the diagnosis as the variable of interest, the second was a model with diagnosis and covariates (age, sex, scanner, comorbidities, and medication), and a third model with the same covariates, but with the interaction between diagnosis and

age as the variable of interest. All models included scanner as a covariate and allowed each site to have their own parameter estimates (slopes and intercepts), thus accommodating between-site differences, further dispensing with the need for explicit harmonization, and avoiding that larger sites would dominate the results. Variance groups [58, 59] were used to allow different variances between GAD and controls. The number of permutations (sign-flippings) was set at 10,000. All results were corrected for multiple testing using the familywise error rate (FWER).

Levene's test for equal variance was used to investigate if participants with GAD have a higher variance in the PAD than HCs. Levene's test was also used to examine the difference across diagnosis and in adults compared to children and adolescents (as preregistered with a threshold of 25 years. This was chosen a priori as an appropriate age for the transition from adolescence to young adulthood). Of note, we do not conduct the exploratory analysis on symptoms scale as described in the preregistration, as we unfortunately did not have enough symptoms scale data at the time of this analysis.

Results

The MAE for the training set was 2.95 years. The MAE for the validation set was 2.30 years. Figure 3 displays a scatterplot of the predicted ages with respect to the true ages in the testing set. Averaged across models in the ensemble, the testing set had an overall MAE of 3.80 years, 2.94 years for control participants, and 6.27 years for subjects with a GAD diagnosis (Supplemental Table 2). A key consideration was to establish the extent to which each model generated consistent predictions and evaluation metrics. Testing results across models revealed stable performance. This was determined by using intraclass correlation (ICC) of the form [2,K] to measure the reliability of the age predictions [60]. To test the reliability of each model's generated age prediction, ICC was calculated across these predictions. The ICC for the predicted age was 0.99, indicating excellent reliability across models. Table

3 displays the results of the MAE and PAD across samples. The MSE for the overall sample was 47.77, whereas the MSE for the control group was 31.72 and for the GAD group 93.57.

Once the average brain age predictions were subtracted from chronological age, the overall PAD was 0.49 years, with a standard deviation of 6.10. For controls, the predicted age difference was 0.36 years (SD = 4.97). For GAD subjects, the PAD was .87 years (SD = 8.53). None of the three GLM models established a statistically significant relationship between PAD and diagnosis (all $p > .05$). This suggests that none of the independent variables (age, sex, scanner, comorbidities, and medication) had any meaningful relationship with the PAD across any of the sites. True age and predicted age significantly correlated ($r = .86, p < .001$), as did true age and PAD ($r = -.26, p < .001$). The greater MAE of subjects with GAD relative to controls was noteworthy.

Levene's test of equality of variances suggests a statistically significant difference in the variance of PAD between controls and subjects with GAD ($W = 442.98, p < .001$), with GAD subjects presenting more variance compared to healthy controls (GAD SD: 8.53, controls SD: 4.97). To examine if age plays a role in the degree of variance between those with and without GAD, the test was implemented after partitioning the data into four groups: healthy volunteers under age 25, individuals with GAD under 25, healthy volunteers over age 25, and individuals with GAD over 25. Subjects with GAD had greater variance, with those over 25 having the greatest variance for all groups. Full results for these tests are shown in Table 4 and 5. A violin plot illustrating the degree of variance in each group is displayed in Figure 4. Mean, median, and standard deviation of the different age and diagnostic groups are shown in Supplemental Table 4.

T tests between the mean ages of individuals with and without GAD were conducted between all sites to examine any significant differences. Several sites were found to have statistically significant differences in age (listed in Table 2). This difference was also significant in the overall sample. Due to these results, we conducted several supplementary analyses, with only the ENIGMA-ANX sites

(excluding ABCD, CMI and BHRC) in an attempt to age match the subjects across diagnostic groups. A T test demonstrated that once these sites were removed, both subjects with and without GAD were age matched. With these sites removed, we reanalyzed the same three GLM models with these new samples. As before, none of these models established a statistically significant relationship between PAD and diagnosis (all $p > .05$). Additionally, we also re-examined the variance in PAD across age and diagnostic groups. The effect of greater variance in subjects with GAD, and greater variance in GAD subjects above 25 was replicated (See supplemental tables 4 and 5).

Discussion

The current study used a CNN to estimate brain age and compute the PAD across individuals with GAD and controls. We used data from 3,500+ healthy individuals across the lifespan to train our model and then tested the model on 6,000+ independent individuals with and without GAD. The difference in PAD between individuals with GAD and controls was then investigated. With respect to our first hypothesis, we did not find a significant difference in PAD between individuals with GAD and controls. However, consistent with our second hypothesis, we found that individuals with GAD have greater PAD variability, particularly in those older than 25 years. After the PAD for each subject was generated, we conducted secondary analyses to examine (1) the relation of diagnostic status across sites, (2) how medication and comorbid disorders relate to PAD, and (3) interaction of diagnosis by age. All three secondary analyses showed no significant relation between the variables of interest and PAD, even after age matching subjects.

Our findings are consistent with prior work showing no significant differences in PAD between GAD and controls and with dimensional measures of anxiety in youth [22]. However, these findings contrast with work within anxiety disorders broadly that found a positive PAD in a substantially smaller sample consisting of participants with GAD, PD and SAD diagnoses, when accounting for medication

effect [18], and other work finding a greater PAD in adolescents with GAD [22]. The heterogeneity of GAD in its clinical presentation, may explain portions of these inconsistent findings.

Although the group difference findings were non-significant, the significantly higher variance in the GAD group is noteworthy, as heterogeneity within psychopathology is a common finding [61, 62]. Other brain age analyses of psychiatric disorders have also found greater variability in the clinical population compared to healthy controls. Han et al. [17] found higher variability among individuals with depression than in healthy individuals. For healthy controls, the MAE was 6.50 and 6.84 years for males and females, respectively. Similarly, within the MDD group, the MAE was 6.72 for males and 7.18 for females. In a replication study, the same authors found that males in the control group had an MAE of 6.85 and females had 7.84. While in the depression group, males showed a larger MAE of 7.57 and females had 9.24, again suggesting greater variability within the clinical population. The same trend appeared in another study examining depression and anxiety disorders [18], where healthy controls manifested an MAE of 5.97, but those in the disorder group had an MAE of 6.73. Finally, Tønnesen et al. [29] found that healthy controls had a MAE of 6.92, patients with bipolar disorder had an MAE = 6.85, and patients with schizophrenia had an MAE = 7.11. This could suggest potentially greater variability in subjects with schizophrenia, but similar variability in bipolar disorder to healthy controls. While notable, these differences in past research were not reported as main findings, nor statistically tested. However, the present study used a statistical procedure to establish a statistically significant difference in variance across those with GAD and controls; this suggests a further avenue for research in brain age prediction for psychiatric disorders. The higher variance in the GAD group indicated that the model performs more poorly in this population. Further investigating the sources of increased variance in the patient population might help understand new mechanisms for the development of psychiatric disorders.

Given that there was a statistically significant age difference in the subjects with GAD, we conducted further supplemental analysis with the non-ENIGMA samples removed to ensure subjects were age matched. This new GLM did not suggest an effect of advanced aging or any other variables having a relationship to the PAD. However, the effect of larger variance in the PAD for individuals with GAD replicated in the age matched sample. This means that the inclusion of sites such as ABCD, where large number of subjects with narrowly distributed ages did not drive the differences in variances between these groups.

The observation of greater PAD variability in individuals with GAD, despite no overall shift in mean brain age, has meaningful implications. Clinically, it underscores the heterogeneity of GAD and raises the possibility that subgroups of patients may experience neurostructural changes that are obscured when averaging across the diagnostic group (e.g., endophenotypes). This possibility, if true, would partially explain the mixed findings in the literature and raises the possibility that PAD variance could represent a transdiagnostic marker of risk or resilience. Moreover, the finding of greater PAD variability in individuals over 25 suggests that brain aging trajectories in GAD may become more divergent in adulthood, highlighting the need for longitudinal research to map the neurostructural progression of the disorder over time. Future studies should explore whether increased PAD variability is associated with symptom chronicity, cognitive decline, treatment response, or comorbidities.

The results of this study should be viewed in the context of several limitations. First, although this study has a large sample size, there was an uneven age distribution across the samples, with a predominance of younger participants largely due to the inclusion of the ABCD data. Additionally, PAD was related to true age, which is a common issue in brain age studies [63, 64]; to address this we included age as a covariate, with no significant results. A second limitation is that, while we constructed our model based on the best existing literature and practices on CNN, this area of research continues to develop quickly. The proposed model does not exhaust possible architectures for brain age prediction, or

even all possibilities among CNN models. It should also be noted that the use of a whole-brain VBM approach in combination with the CNN architecture does not allow for group comparisons of structural changes in specific regions of interest (ROI) previously associated with GAD, as is common practice in the neuroimaging literature. However, the CNN approach demonstrates more accurate overall brain age estimations compared to linear ROI approaches in healthy subjects [27]. Thus, the CNN architecture is more sensitive to subject level spatial information making it a better model choice for the proposed analyses with the downside of lacking the granularity of an ROI-based linear approach.

In summary, we found no significant difference in PAD between individuals with GAD compared to healthy controls. This is consistent with prior work in adolescents with anxiety symptoms that found no association between PAD and anxiety. We found PAD was not associated with medication use or comorbid disorders related to PAD, and found no interaction between age and diagnostic status with PAD. In other words, the results in this analysis do not support the hypothesis that GAD is associated with accelerated brain aging. Of note, we did find significantly larger variability in PAD for GAD compared to the healthy controls. This difference in variability between a psychiatric group and controls is an area that should be more robustly examined and reported in future studies of PAD. Lastly, the training and testing samples assembled for this study are substantially larger than previous work within anxiety domains and brain PAD. Reporting these null results from a large mega-analysis is important as it may illustrate that a more detailed approach is necessary to fully capture the heterogeneity underlying GAD. It might be that specific subpopulations, treatment exposures, or specific symptoms are associated with differences in brain age; however, it is difficult to disentangle these without harmonized data collection.

474 Data and code sharing

475 The data analyzed in this study is a combination of open datasets and the collaborative work of
476 ENIGMA-Anxiety. The ENIGMA-Anxiety Working Group accepts research proposals from *bona fide*
477 researchers. The data are shared by each participating PI after approval of the project. For more
478 information please see: <https://enigma.ini.usc.edu/ongoing/enigma-anxiety/>. The code used in the
479 analysis described here will be available at
480 https://github.com/coreyjr2/ENIGMA_BrainAge_GAD_structural.

481

482

483

Conflicts of interest

PMP has received speakers' honoraria from Daiichi Sankyo, Libbs, Sandoz, Novartis, and Abbot. NJ is MPI of a research related grant from Biogen Inc, for work unrelated to this manuscript. JCS is a consultant for PFIZER, SUNOVIAN, SANOFI, JOHNSON & JOHNSON, Livanova, and Boehringer Ingelheim. Material support (for IOL site, led by MA and GJD) was provided by Neuronetics. FA is Section Editor of NeuroImage: Clinical; has received speaker honoraria from Philips, Novartis, and Biogen Idec; and receives or has received research supports from the Italian Ministry of Health, AriSLA (Fondazione Italiana di Ricerca per la SLA), and the European Research Council. EC has received research supports from the Italian Ministry of Health. MF is Editor-in-Chief of the Journal of Neurology; received compensation for consulting services and/or speaking activities from Biogen Idec, Merck-Serono, Novartis, and Teva Pharmaceutical Industries; and receives research support from Biogen Idec, Merck-Serono, Novartis, Teva Pharmaceutical Industries, Roche, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and ARiSLA (Fondazione Italiana di Ricerca per la SLA). MF is Editor-in-Chief of the Journal of Neurology; received compensation for consulting services and/or speaking activities from Biogen Idec, Merck-Serono, Novartis, and Teva Pharmaceutical Industries; and receives research support from Biogen Idec, Merck-Serono, Novartis, Teva Pharmaceutical Industries, Roche, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and ARiSLA (Fondazione Italiana di Ricerca per la SLA). GAF owns equity in Alto Neuroscience and has served as a consultant for SynpaseBio AI and Alto Neuroscience. MPP is an advisor to Spring Care, Inc., a behavioral health startup, he has received royalties for an article about methamphetamine in UpToDate. MBS has in the past 3 years been a consultant for Actelion, Acadia Pharmaceuticals, Aptinyx, Bionomics, BioXcel Therapeutics, Clexio, EmpowerPharm, Epivario, GW Pharmaceuticals, Janssen, Jazz Pharmaceuticals, Roche/Genentech, and Oxeia Biopharmaceuticals; and has stock options in Oxeia Biopharmaceuticals and Epivario. DJS has received consultancy honoraria from Discovery Vitality, Johnson & Johnson, Kanna, L'Oreal, Lundbeck, Orion, Sanofi, Servier, Takeda and Vistagen. K.H. is a scientific advisor and received virtual stock options of Mental Tech GmbH, which develops an AI-based chatbot providing mental health support.

525 **Table 1**

526 *Number of Subjects and Descriptive Statistics - Training Sample*

Site	Total n	Average Age	Standard Deviation	Age Min	Age Max
HCP	412	28.71	3.67	22	37
HCP Development	612	14.52	4.04	5.58	21.92
OASIS	334	68.93	9.29	43	93
PING	588	12.42	5.12	3.08	21.67
PNC	1,565	15.02	3.66	8.08	23.08
Total/Average	3,511	21.23	16.85	16.35	39.33

527 HCP: Human Connectome Project, HCP Development Human Connectome Project-
528 Development, OASIS: Open Access Series of Imaging Studies, PING: Pediatric Imaging,
529 Neurocognition, and Genetics study. PNC: Philadelphia Neurodevelopmental Cohort
530

531
532
533
534
535
536
537
538
539

540 **Table 2**541 *Number of Participants and Descriptive Statistics - Testing Sample*

Site	Total n	Age Min	Age Max	HC Sample	GAD Sample	GAD Average Age (SD)	HC Average Age (SD)	<i>p</i>
ABCD	2,857	8.91	11	2,778	79	10.05 (.62)	9.92 (.62)	.07
Barcelona	69	18	40	46	23	23.61 (4.3)	21.71 (4.48)	.10
Baylor	226	12	79	128	98	28.84 (9.18)	24.39 (12.84)	<.05
BHRC	767	7.02	21.37	458	309	13.38 (3.15)	11.70 (2.64)	<.05
Boystown	95	13	18	45	50	15.82 (1.35)	15.44 (1.62)	.23
Chicago-Milad	43	18	59	16	27	30.25 (9.71)	32.37 (12.96)	.58
Chicago-Phan	141	18	60	41	100	26.36 (8.51)	25.29 (9.75)	.54
Cincinnati	20	12	18	11	9	13.66 (2.05)	14.18 (1.89)	.59
CMI-HBN	185	5.01	21.36	97	88	12.24 (3.53)	9.57 (3.21)	<.05
Dresden	93	18	51	46	47	30.23 (9.94)	28.47 (7.90)	.35
Duke	40	6	10	19	21	7 (1.06)	7.47 (.99)	.16
SHIP	32	41	70	22	10	55.9 (7.95)	57.72 (8.01)	.57

Harvard	237	18	40	51	186	24.64 (4.31)	24.68 (3.77)	.95
Houston	263	8	68.23	251	12	16.69 (9.02)	29.21 (14.39)	<.05
Milan	91	20.89	72.57	61	30	44.31 (14.91)	34.11 (12.42)	<.05
Muenster	53	19	56	29	24	29.45 (10.01)	27.82 (9.04)	.54
Pittsburgh- Andreescu	64	19	82	26	38	57.78 (16.51)	56.84 (20.22)	.84
Pittsburgh-Price	69	18	54	0	69	N/A	N/A	N/A
PROTAIA	43	13	22	25	18	16.67 (2.26)	17.80 (2.54)	.14
Rome	38	18	55	20	18	29.55 (6.92)	29.15 (9.18)	.88
San Raffaele	89	23	63	69	20	44.65 (10.61)	45.26 (10.97)	.83
SDAN	295	8.1	51.14	156	139	12.77 (2.78)	13.70 (4.37)	<.05
SNFA	60	19	50	38	22	31.27 (9.97)	29 (7.17)	.36
Stony Brook	60	18	49	19	41	22.87 (5.91)	21.57 (5.74)	.43
UCSD	91	17	53	47	44	28.36 (10.75)	23.40 (8.27)	<.05
Wash U	63	8	12	32	31	9.96 (1.44)	9.90 (1.15)	.85
IOL	63	19	71	21	42	39.92 (13.38)	41.19 (16.17)	.76

Total/Average	6,147	15.66	46.58	4,552	1,595	22.16 (13)	14.53 (10.59)	<.05
---------------	-------	-------	-------	-------	-------	------------	------------------	------

542 GAD: generalized anxiety disorder, HC: healthy controls, p: p value, ABCD: Adolescent Brain
543 Cognitive Development Study, BHRCS: Brazilian High Risk Cohort Study, CMI-HBN: Child Mind
544 Institute Healthy Brain Network, IOL: Institute of Living, HCP: Human Connectome Project, HCPdev:
545 Human Connectome Project Development, PING: Pediatric Imaging, Neurocognition, and Genetics
546 Study, OASIS: Open Access Series of Imaging Studies, PNC: Philadelphia Neurodevelopmental Cohort,
547 PROTAIA: Anxiety Disorders Program for Child and Adolescent Psychiatry, SDAN: Section on
548 Development and Affective Neuroscience, SNFA: Section on Neurobiology of Fear and Anxiety,
549 UCSD: University of California – San Diego, WashU: Washington University. SHIP: Study of Health in
550 Pomerania.

551 **Table 3**552 *Testing Set Age, Predicted Age, and PAD Organized by Sample*

	GAD n	GAD PAD Mean	GAD PAD SD	HC n	HC PAD Mean	HC PAD SD
ABCD	79	0.179	1.18	2,778	0.58	1.19
Barcelona	23	-1.45	6.03	46	-1.83	4.79
Baylor	98	5.36	7.71	128	1.06	5.40
BHRC	309	1.60	4.92	458	0.85	4.53
Boystown	50	-2.21	2.18	45	-2.92	1.72
Chicago-Milad	27	4.31	8.52	16	4.89	8.97
Chicago-Phan	100	.94	5.95	41	1.11	7.38
Cincinnati	9	-1.14	0.96	11	-0.88	1.29
CMI	88	2.22	4.03	97	1.20	2.32
Dresden	47	9.23	7.01	46	10.92	7.41
Duke	21	0.23	1.33	19	-0.48	0.91
SHIP	10	6.25	6.03	22	1.69	7.17
Harvard	186	4.76	6.09	51	4.43	5.36
Houston	12	-2.67	5.44	251	-4.61	8.16

Milan	30	-7.04	9.98	61	-8.46	7.63
Muenster	24	12.1	8.72	29	9.04	8.96
Pittsburgh- Andreescu	38	-23.27	11.46	26	-19.83	12.25
Pittsburgh- Price	69	-8.51	5.82	N/A	N/A	N/A
PROTAIA	18	6.51	5.90	25	8.72	8.64
Rome	18	5.24	7.49	20	3.72	7.63
San Raffaele	20	11.65	4.99	69	5.41	8.72
SDAN	139	-0.19	1.42	156	0.14	2.41
SNFA	22	1.58	8.56	38	-0.81	9.93
Stony Brook	41	9.26	7.46	19	12.76	6.74
UCSD	44	-3.09	9.28	47	-0.99	6.33
Wash U	31	2.42	1.13	32	2.47	1.21
IOL	42	-14.80	8.64	21	-13.17	9.31

PAD predicted age difference, SD standard deviation, GAD generalized anxiety disorder, HC healthy controls, ABCD: Adolescent Brain Cognitive Development Study, BHRCS: Brazilian High Risk Cohort Study, CMI-HBN: Child Mind Institute Healthy Brain Network, IOL: Institute of Living, HCP: Human Connectome Project, HCPdev: Human Connectome Project Development, PING: Pediatric Imaging, Neurocognition, and Genetics Study, OASIS: Open Access Series of Imaging Studies, PNC: Philadelphia Neurodevelopmental Cohort, PROTAIA: Anxiety Disorders Program for Child and Adolescent Psychiatry, SDAN: Section on Development and Affective Neuroscience, SNFA: Section on Neurobiology of Fear and Anxiety, UCSD: University of California – San Diego, WashU: Washington University. SHIP: Study of Health in Pomerania

553
554
555
556
557
558
559
560
561
562

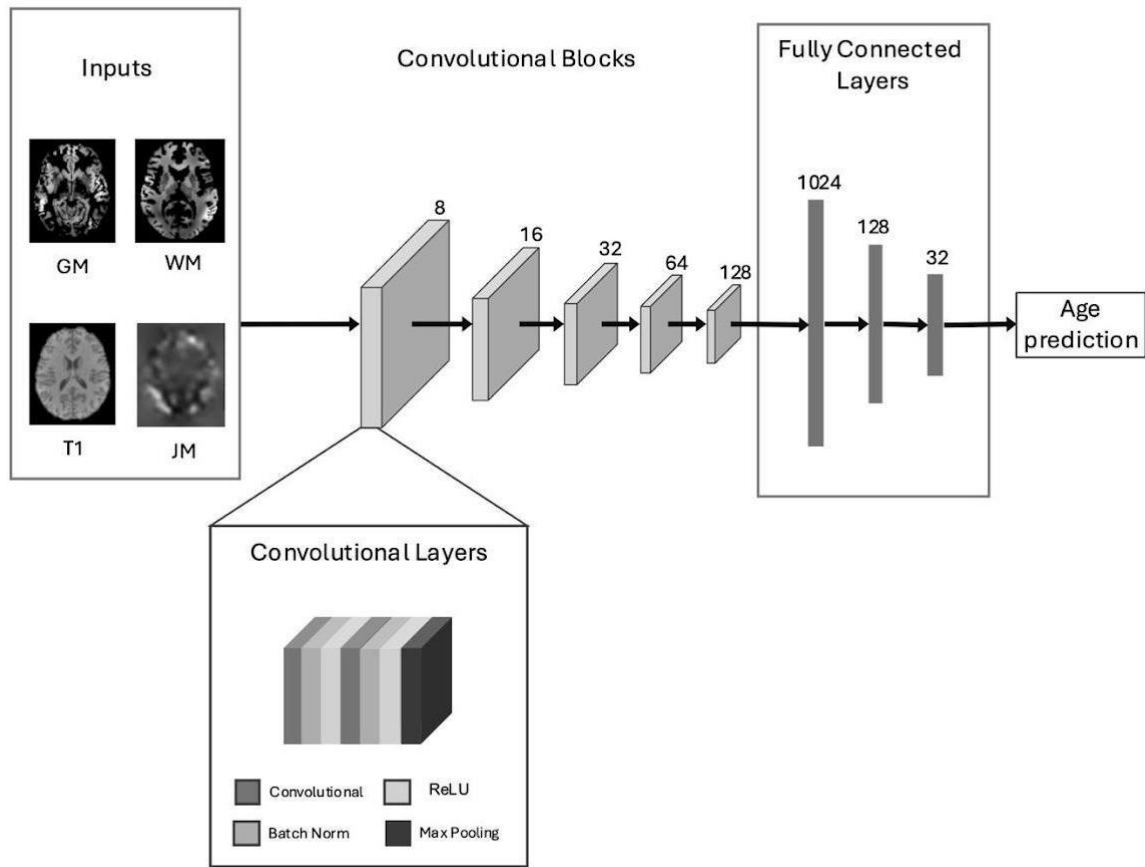
Table 4*Characteristics of PAD by Diagnosis and Age*

Group	Mean PAD	Median PAD	Standard Deviation	Variance
GAD < 25	1.86	0.42	5.25	27.64
HC < 25	0.77	0.51	3.01	9.06
GAD > 25	-1.42	0.10	13.08	171.136
HC > 25	-2.86	-2.70	11.57	134.04

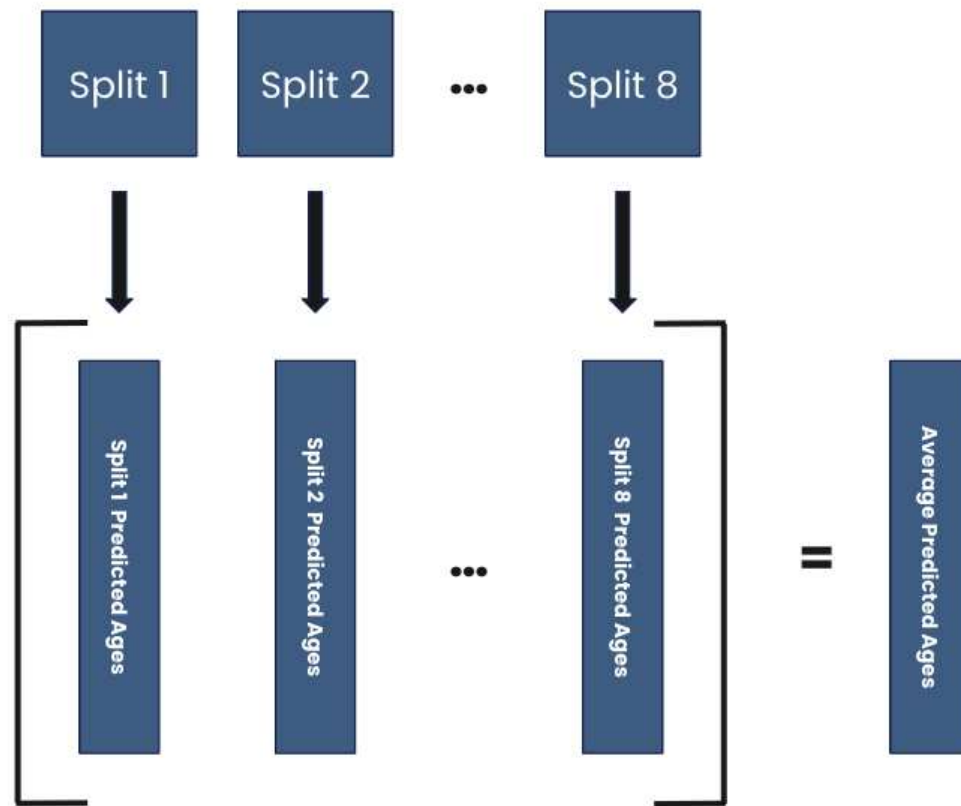
Table 5*Levene's Tests Comparing Variances in PAD across Diagnostic and Age Groups*

Group 1	Group 2	W	P
GAD < 25	HC < 25	348.52	<.001
GAD < 25	GAD > 25	489.6	<.001
GAD < 25	HC > 25	394.01	<.001
HC < 25	GAD > 25	2541.49	<.001
HC < 25	HC > 25	2221.09	<.001
GAD > 25	HC > 25	6.74	<.001

Figure 1
Overview of the CNN Model Architecture.



Note. The CNN architecture was designed with four input channels for gray and white matter segmentations, as well as the T1-weighted and Jacobian images. The model consists of five repeated convolutional blocks, each containing 3D convolutional layers, batch normalization, ReLU activations and max pooling. The number of filters used in each block are denoted above. After the last convolutional block, the model includes a flattened layer.

Figure 2*Model Ensemble Procedure*

Note. Procedure for generating average predictions for the model ensemble. Eight splits of the training data were created, which used a different 90/10% split of training to validation data. Each of these splits was given to a unique instantiation of the model with identical architecture but instantiated with random weights.

Figure 3
Scatter Plot of True Age vs. Predicted Age Sorted By Diagnostic Status

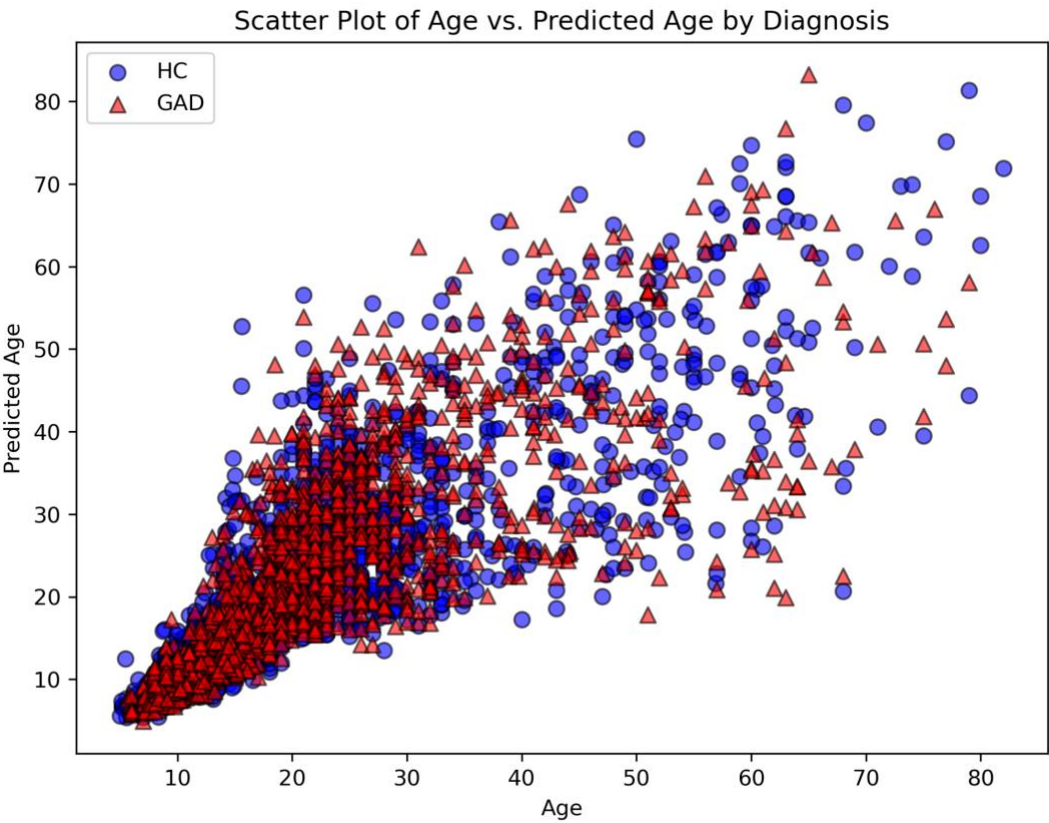
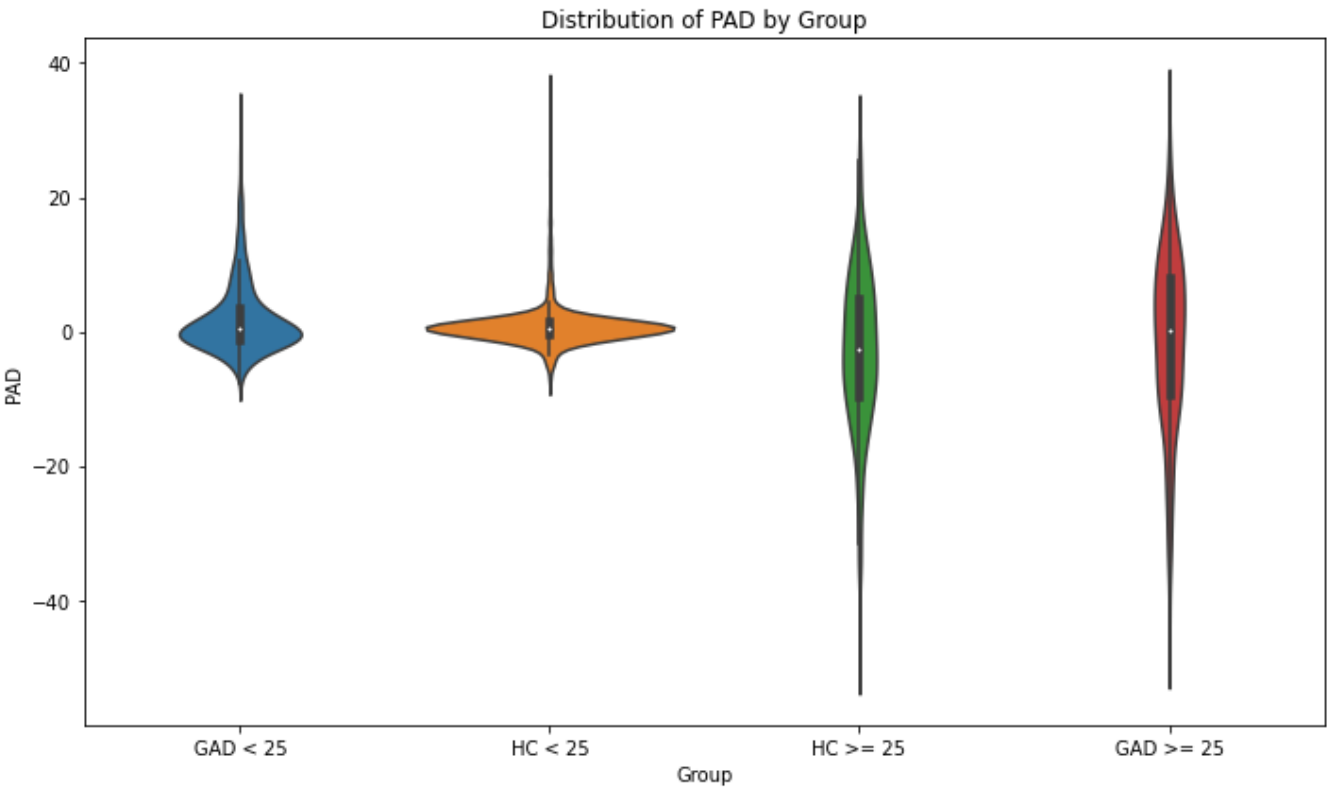


Figure 4
Violin Plot of Variance in PAD across Age and Diagnostic Groups



660 **Supplementary Material**

661 **Supplemental Table 1**

662 *Sample Selection Procedure*

Site	Subjects with VBM	After data entry exclusions	After QC (Final Sample)
ABCD	2,994	2,994	2,857
Barcelona	91	91	69
Baylor	228	228	226
BHRC	799	797	767
Boystown	95	95	95
Chicago-Milad	43	43	43
Chicago-Phan	147	147	141
Cincinnati	20	20	20
CMI	198	198	185
Dresden	94	94	93
Duke	40	40	40
SHIP	36	36	32
Harvard	260	260	237
HCP	414	414	412
HCPdev	652	652	612

Houston	273	273	263
Milan	97	97	91
Muenster	54	54	53
OASIS	480	480	334
PING	741	729	588
Pittsburgh- Andreescu	79	79	64
Pittsburgh- Price	69	69	69
PNC	1,597	1,597	1,565
PROTAIA	44	44	43
Rome	40	40	38
San Raffaele	92	92	89
SDAN	295	295	295
SNFA	63	63	60
Stony Brook	61	61	60
UCSD	96	96	91
Wash U	63	63	63
IOL	64	64	63
Total	10,319	10,305	9,658

VBM: Voxel Based Morphometry, QC: quality control, GAD: generalized anxiety disorder, HC: healthy controls, ABCD: Adolescent Brain Cognitive Development Study, BHRCS: Brazilian High Risk

Cohort Study, CMI-HBN: Child Mind Institute Healthy Brain Network, IOL: Institute of Living, HCP: Human Connectome Project, HCPdev: Human Connectome Project Development, PING: Pediatric Imaging, Neurocognition, and Genetics Study, OASIS: Open Access Series of Imaging Studies, PNC: Philadelphia Neurodevelopmental Cohort, PROTAIA: Anxiety Disorders Program for Child and Adolescent Psychiatry, SDAN: Section on Development and Affective Neuroscience, SNFA: Section on Neurobiology of Fear and Anxiety, UCSD: University of California – San Diego, WashU: Washington University. SHIP: Study of Health in Pomerania

711 **Supplemental Table 2**

712 *Model Performance by Each Model in the Ensemble*

	Total Test MAE	HC Test MAE	GAD Test MAE
Model 1	3.70	2.85	5.97
Model 2	3.57	2.80	6.46
Model 3	4.25	3.22	5.61
Model 4	3.33	2.66	5.66
Model 5	3.28	3.29	6.63
Model 6	4.31	3.29	6.45
Model 7	4.14	3.06	5.94
Model 8	3.82	2.91	5.87
Overall	3.80	2.94	6.27

713 MAE: mean absolute error. HC: healthy controls, GAD: generalized anxiety disorder.

714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733

734 **Supplemental Table 3**

735 *Overview of variables and contrasts of interest in the secondary analysis.*

Dependent Variable		Independent Variables		
Predicted Age Difference (PAD)	Variables of interest	Model 1 (group comparison)	Model 2 (Covariates)	Model 3 (Moderation)
		Diagnosis	Diagnosis	Diagnosis
	Nuisance Variables		Age Age ² Scanner Medication Comorbid SAD, PD, AG, SPH, other anxiety disorders, MDD, OCD, PTSD, SUD, other disorders	Age Age ² Scanner Medication Comorbid SAD, PD, AG, SPH, other anxiety disorders, MDD, OCD, PTSD, SUD, other disorders Age * Diagnosis

736 PAD: predicted age difference, SAD: separation anxiety disorder, PD: panic disorder, AG: agoraphobia,
 737 SPH: specific phobia, MDD: major depressive disorder, OCD: obsessive compulsive disorder, PTSD:
 738 post-traumatic stress disorder, SUD: substance use disorder

740

741

742

Supplemental Table 4

Descriptive statistics of predicted age difference by age and anxiety disorder status with ABCD, CMI, and BHRC removed.

Group	Mean	Median	Standard Deviation	Variance
GAD < 25	2.14	0.61	5.80	33.67
HC < 25	1.45	0.13	5.64	31.85
GAD > 25	-1.42	-2.70	11.57	171.13
HC > 25	-2.86	-2.70	11.57	134.04

Supplemental Table 5

Levene's Tests Comparing Variances in PAD across Diagnostic and Age Groups with ABCD, CMI, and BHRC removed

Group 1	Group 2	W	P
GAD < 25	HC < 25	10.71	<.001
GAD < 25	GAD > 25	436.01	<.001
GAD < 25	HC > 25	346.86	<.001
HC < 25	GAD > 25	586.32	<.001
HC < 25	HC > 25	481.04	<.001
GAD > 25	HC > 25	6.74	<.001

References

1. Szuhany, K.L. and N.M. Simon, *Anxiety Disorders: A Review*. JAMA, 2022. 328(24): p. 2431-2445.
2. Wittchen, H.-U., *Generalized anxiety disorder: prevalence, burden, and cost to society*. Depression and Anxiety, 2002. 16(4): p. 162-171.
3. Kessler, R.C., et al., *Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication*. Archives of General Psychiatry, 2005. 62(6): p. 617.
4. Kessler, R.C., et al., *Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication*. Archives of General Psychiatry, 2005. 62(6): p. 593.
5. Cole, J.H., et al., *Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker*. NeuroImage, 2017. 163: p. 115-124.
6. Smith, S.M., et al., *Estimation of brain age delta from brain imaging*. NeuroImage, 2019. 200: p. 528-539.
7. Hilbert, K., U. Lueken, and K. Beesdo-Baum, *Neural structures, functioning and connectivity in Generalized Anxiety Disorder and interaction with neuroendocrine systems: A systematic review*. Journal of Affective Disorders, 2014. 158: p. 114-126.
8. Brambilla, P., et al., *White-matter abnormalities in the right posterior hemisphere in generalized anxiety disorder: a diffusion imaging study*. Psychological Medicine, 2011. 42(2): p. 427-434.
9. Andreescu, C., et al., *The ages of anxiety: differences across the lifespan in the default mode network functional connectivity in generalized anxiety disorder*. International Journal of Geriatric Psychiatry, 2013. 29(7): p. 704-712.

- 809 10. Sylvester, C.M., et al., *Functional network dysfunction in anxiety and anxiety disorders*. Trends
810 in Neurosciences, 2012. 35(9): p. 527-535.
- 811 11. Franke, K., et al., *Brain maturation: Predicting individual BrainAGE in children and*
812 *adolescents using structural MRI*. NeuroImage, 2012. 63(3): p. 1305-1312.
- 813 12. Gaser, C., et al., *BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to*
814 *Alzheimer's Disease*. PLoS ONE, 2013. 8(6): p. e67346.
- 815 13. Cole, J.H., R. Leech, and D.J. Sharp, *Prediction of brain age suggests accelerated atrophy after*
816 *traumatic brain injury*. Annals of Neurology, 2015. 77(4): p. 571-581.
- 817 14. Koutsouleris, N., et al., *Accelerated Brain Aging in Schizophrenia and Beyond: A*
818 *Neuroanatomical Marker of Psychiatric Disorders*. Schizophrenia Bulletin, 2013. 40(5): p.
819 1140-1153.
- 820 15. Schnack, H.G., et al., *Accelerated Brain Aging in Schizophrenia: A Longitudinal Pattern*
821 *Recognition Study*. American Journal of Psychiatry, 2016. 173(6): p. 607-616.
- 822 16. Keding, T.J., et al., *Differential Patterns of Delayed Emotion Circuit Maturation in Abused Girls*
823 *With and Without Internalizing Psychopathology*. American Journal of Psychiatry, 2021.
824 178(11): p. 1026-1036.
- 825 17. Han, L.K.M., et al., *Brain aging in major depressive disorder: results from the ENIGMA major*
826 *depressive disorder working group*. Molecular Psychiatry, 2020. 26(9): p. 5124-5139.
- 827 18. Han, L.K.M., et al., *Contributing factors to advanced brain aging in depression and anxiety*
828 *disorders*. Translational Psychiatry, 2021. 11(1).
- 829 19. Kaufmann, T., et al., *Common brain disorders are associated with heritable patterns of apparent*
830 *aging of the brain*. Nature Neuroscience, 2019. 22(10): p. 1617-1623.

- 831 20. Blake, K.V., et al., *Advanced brain ageing in adult psychopathology: A systematic review and*
832 *meta-analysis of structural MRI studies*. Journal of Psychiatric Research, 2023. 157: p. 180-191.
- 833 21. Cropley, V.L., et al., *Brain-Predicted Age Associates With Psychopathology Dimensions in*
834 *Youths*. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2021. 6(4): p. 410-
835 419.
- 836 22. Zhou, Z., et al., *Differential effects of generalized anxiety and separation anxiety on brain*
837 *structural development during adolescence*. Journal of Affective Disorders, 2023. 339: p. 478-
838 485.
- 839 23. Lindsay, G.W., *Convolutional Neural Networks as a Model of the Visual System: Past, Present,*
840 *and Future*. Journal of Cognitive Neuroscience, 2021. 33(10): p. 2017-2031.
- 841 24. Mendes, S.L., et al., *Estimating Gender and Age from Brain Structural MRI of Children and*
842 *Adolescents: A 3D Convolutional Neural Network Multitask Learning Model*. Computational
843 Intelligence and Neuroscience, 2021. 2021(1).
- 844 25. Christman, S., et al., *Accelerated brain aging predicts impaired cognitive performance and*
845 *greater disability in geriatric but not midlife adult depression*. Translational Psychiatry, 2020.
846 10(1).
- 847 26. Bashyam, V.M., et al., *MRI signatures of brain age and disease over the lifespan based on a*
848 *deep brain network and 14 468 individuals worldwide*. Brain, 2020. 143(7): p. 2312-2324.
- 849 27. Ueda, M., et al., *An Age Estimation Method Using 3D-CNN From Brain MRI Images*, in *2019*
850 *IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. 2019, IEEE. p. 380-
851 383.
- 852 28. Leonardsen, E.H., et al., *Deep neural networks learn general and clinically relevant*
853 *representations of the ageing brain*. NeuroImage, 2022. 256: p. 119210.

- 854 29. Tønnesen, S., et al., *Brain Age Prediction Reveals Aberrant Brain White Matter in Schizophrenia*
855 *and Bipolar Disorder: A Multisample Diffusion Tensor Imaging Study*. Biological Psychiatry:
856 Cognitive Neuroscience and Neuroimaging, 2020. 5(12): p. 1095-1103.
- 857 30. Ge, R., et al., *Normative modelling of brain morphometry across the lifespan with CentileBrain:*
858 *algorithm benchmarking and model optimisation*. Lancet Digit Health, 2024. 6(3): p. e211-e221.
- 859 31. Bethlehem, R.A.I., et al., *Brain charts for the human lifespan*. Nature, 2022. 604(7906): p. 525-
860 533.
- 861 32. Bas-Hoogendam, J.M., et al., *ENIGMA-Anxiety Working Group - Generalized Anxiety Disorder*
862 *Brain Age Project*. 2023, OSF Registries.
- 863 33. Harrewijn, A., et al., *Cortical and subcortical brain structure in generalized anxiety disorder:*
864 *findings from 28 research sites in the ENIGMA-Anxiety Working Group*. Translational
865 Psychiatry, 2021. 11(1).
- 866 34. Jonsson, B.A., et al., *Brain age prediction using deep learning uncovers associated sequence*
867 *variants*. Nature Communications, 2019. 10(1): p. 5409.
- 868 35. Zugman, A., et al., *Mega-analysis methods in ENIGMA: The experience of the generalized*
869 *anxiety disorder working group*. Human Brain Mapping, 2020. 43(1): p. 255-277.
- 870 36. Van Essen, D.C., et al., *The WU-Minn Human Connectome Project: An overview*. NeuroImage,
871 2013. 80: p. 62-79.
- 872 37. LaMontagne, P.J., et al., *OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset*
873 *for Normal Aging and Alzheimer Disease*. 2019, Cold Spring Harbor Laboratory.
- 874 38. Satterthwaite, T.D., et al., *The Philadelphia Neurodevelopmental Cohort: A publicly available*
875 *resource for the study of normal and abnormal brain development in youth*. Neuroimage, 2016.
876 124(Pt B): p. 1115-1119.

- 877 39. Jernigan, T.L., et al., *The Pediatric Imaging, Neurocognition, and Genetics (PING) Data*
878 *Repository*. NeuroImage, 2016. 124: p. 1149-1154.
- 879 40. Casey, B.J., et al., *The Adolescent Brain Cognitive Development (ABCD) study: Imaging*
880 *acquisition across 21 sites*. Dev Cogn Neurosci, 2018. 32: p. 43-54.
- 881 41. Alexander, L.M., et al., *An open resource for transdiagnostic research in pediatric mental health*
882 *and learning disorders*. Scientific Data, 2017. 4(1): p. 170181.
- 883 42. Salum, G.A., et al., *High risk cohort study for psychiatric disorders in childhood: rationale,*
884 *design, methods and preliminary results*. Int J Methods Psychiatr Res, 2015. 24(1): p. 58-73.
- 885 43. Douaud, G., et al., *Anatomically related grey and white matter abnormalities in adolescent-onset*
886 *schizophrenia*. Brain, 2007. 130(9): p. 2375-2386.
- 887 44. Ashburner, J. and K.J. Friston, *Voxel-Based Morphometry: The Methods*. NeuroImage, 2000.
888 11(6): p. 805-821.
- 889 45. Lu, B., et al., *A practical Alzheimer's disease classifier via brain imaging-based deep learning*
890 *on 85,721 samples*. Journal of Big Data, 2022. 9(1): p. 101.
- 891 46. Good, C.D., et al., *A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human*
892 *Brains*. NeuroImage, 2001. 14(1): p. 21-36.
- 893 47. Fonov, V., et al., *Unbiased average age-appropriate atlases for pediatric studies*. NeuroImage,
894 2011. 54(1): p. 313-327.
- 895 48. Fortin, J.-P., et al., *Harmonization of cortical thickness measurements across scanners and sites*.
896 NeuroImage, 2018. 167: p. 104-120.
- 897 49. Yu, Y., et al., *Brain-age prediction: Systematic evaluation of site effects, and sample age range*
898 *and size*. Human Brain Mapping, 2024. 45(10).
- 899 50. Harris, C.R., et al., *Array programming with NumPy*. Nature, 2020. 585(7825): p. 357-362.

- 900 51. Paszke, A., et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*.
901 2019, arXiv.
- 902 52. Srivastava, N., et al., *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*.
903 Journal of Machine Learning Research, 2014. 15(56): p. 1929-1958.
- 904 53. Kingma, D.P. and J. Ba, *Adam: A Method for Stochastic Optimization*. 2014, arXiv.
- 905 54. Pascanu, R., T. Mikolov, and Y. Bengio, *On the difficulty of training Recurrent Neural*
906 *Networks*. 2012, arXiv.
- 907 55. Ganaie, M.A., et al., *Ensemble deep learning: A review*. Engineering Applications of Artificial
908 Intelligence, 2022. 115: p. 105151.
- 909 56. Khosla, M., et al., *Ensemble learning with 3D convolutional neural networks for functional*
910 *connectome-based prediction*. NeuroImage, 2019. 199: p. 651-662.
- 911 57. Kuo, C.-Y., et al., *Improving Individual Brain Age Prediction Using an Ensemble Deep*
912 *Learning Framework*. Frontiers in Psychiatry, 2021. 12.
- 913 58. Winkler, A.M., et al., *Permutation inference for the general linear model*. NeuroImage, 2014.
914 92: p. 381-397.
- 915 59. Winkler, A.M., et al., *Multi-level block permutation*. Neuroimage, 2015. 123: p. 253-68.
- 916 60. Koo, T.K. and M.Y. Li, *A Guideline of Selecting and Reporting Intraclass Correlation*
917 *Coefficients for Reliability Research*. Journal of Chiropractic Medicine, 2016. 15(2): p. 155-163.
- 918 61. Feczko, E. and D.A. Fair, *Methods and Challenges for Assessing Heterogeneity*. Biol Psychiatry,
919 2020. 88(1): p. 9-17.
- 920 62. Olshan, S.M., et al., *Making the most of errors: Utilizing erroneous classifications generated by*
921 *machine-learning models of neuroimaging data to capture disorder heterogeneity*. Journal of
922 Psychopathology and Clinical Science, 2024. 133(8): p. 678-689.

- 923 63. Liang, H., F. Zhang, and X. Niu, *Investigating systematic bias in brain age estimation with*
924 *application to post-traumatic stress disorders*. Human Brain Mapping, 2019. 40(11): p. 3143-
925 3152.
- 926 64. More, S., et al., *Brain-age prediction: A systematic comparison of machine learning workflows*.
927 NeuroImage, 2023. 270: p. 119947.
928